PROPERTIES OF REGRESSION ESTIMATES BASED ON CENSORED SURVIVAL DATA

in A Festschrift for Erich L. Lehmann (1982).
ed. by P.J. Bickel, K.A. Doksum, and J.L. Hodges, Jr.

Kjell A. Doksum
University of California, Berkeley

Brian S. Yandell
University of Wisconsin, Madison

ABSTRACT

This paper gives consistency results, asymptotic distributions, and confidence bands for nonparametric regression curves based on data subject to random right censoring. Two of the nonparametric estimates, Beran's running product limit median and a running product limit mean, are compared with the Cox regression estimate and a least squares type estimate using the Stanford heart transplant data.

1. INTRODUCTION

We consider experiments in which a response variable T, such as survival time or failure time, is subject to random right censoring and has a distribution depending on covariates $Z_1,\dots,Z_p$, such as age, sex, and so on. The object is to investigate how the distribution of T is influenced by the covariates $Z_1,\dots,Z_p$. Cox (1972, 1975) did this by letting the power parameter Δ in Lehmann's (1953) model $1-(1-F)^{\Delta}$ be a known function of $\sum_{j=1}^{p} \beta_j Z_j$, where $\beta_1,\dots,\beta_p$ are regression parameters. Miller (1976); Buckley and James (1979); and Koul, Susarla, and van Ryzin (1981) used a log-linear model involving $p+1$ linear regression

---

This research was supported in part by National Science Foundation Grant MCS-8102349.

$$g'(x) = \begin{cases} -\frac{1}{2} e^{-2} \pi x \sin \pi x, & 0 < x \le 2 \\ xe^{-x}(2 - x), & x \ge 2. \end{cases}$$

Therefore,

$$\int_2^\infty [g'(y)/y]dy = \int_2^\infty e^{-y}(2 - y)dy = -e^{-2}$$

and

$$\int_1^2 [g'(y)/y]dy = -\frac{1}{2} e^{-2}\pi \int_1^2 \sin \pi y \, dy = e^{-2}.$$

Since $g'(x) < 0$ for $0 < x < 1$, it follows that g satisfies (3.10).

REFERENCES

HODGES, J. L., and LEHMANN, E.L. (1954), "Matching in Paired Comparisons," *Annals of Mathematical Statistics*, 25, 787-791.

OLSHEN, R.A., and SAVAGE, L.J. (1970), "A Generalized Unimodality," *Journal of Applied Probability*, 7, 21-34.

type parameters to model the dependence between T and $Z_1,\ldots,Z_p$.

Recently, Beran (1981) introduced and proved strong consistency of nonparametric estimates of the conditional survival distribution $S(t|\mathbf{z}) = P\{T \geq t|\mathbf{Z} = \mathbf{z}\}$ and of the median $m(\mathbf{z})$ of this distribution, where $\mathbf{Z} = (Z_1,\ldots,Z_p)$ is the vector of covariates.

In this paper, we consider estimates of the nonparametric regression surface $\mu(\mathbf{z}) = E(T|\mathbf{Z} = \mathbf{z})$ under conditions where this function is identifiable. We give conditions under which the estimates are consistent and, in the case of one covariate z and censoring independent of that covariate, derive the asymptotic distribution of $\sup_z |\hat{\mu}(z) - \hat{\mu}(z)|$ appropriately standardized and give simultaneous confidence bands for $\mu(z)$. These bands are especially useful since they give an indication of how reliable the estimates of $\mu(z)$ are as well as make it possible to test whether a log-linear model or Cox-type model is correct.

The conditional probabilities we deal with are determined up to a null set of $\mathbf{z}$s. In order to avoid writing "almost all $\mathbf{z}$" throughout, we assume that the $\mathbf{z}$s we consider are in the complement of this null set.

2. CONSISTENT ESTIMATES

2.1 The General Case

We suppose that we have available a random sample of n subjects with the ith subject having failure time T_i and covariates $Z_{i1},\ldots,Z_{ip}$. In experiments with censoring, not all the T_i are observed. Instead we observe $Y_i = \min(T_i,C_i)$ and $J_i = I[T_i \leq C_i]$, where $C_1,\ldots,C_n$ are censoring variables, and I is the indicator function. It is assumed that $\{(Y_i,J_i,\mathbf{Z}_i),$ $i = 1,\ldots,n\}$ are i.i.d. with the same distribution as $(Y,J,\mathbf{Z})$, where $Y = \min(T,C)$, $J = I[T \leq C]$, and $\mathbf{Z}_i = (Z_{i1},\ldots,Z_{ip})$. Moreover, it is assumed that T and C are conditionally independent given $\mathbf{Z}$.

Beran's (1981) running product limit estimate (the BRPLE) of the conditional survival function $S(t|\mathbf{z})$ is

$$S_{n,k}(t|\mathbf{z}) = \prod_{\{j|t_j<t\}} [1 - \hat{\lambda}_{j,k}(\mathbf{z})],$$

where $t_1 < \ldots < t_r$, $r \leq n$, are the ordered distinct observed failure times,

$$\hat{\lambda}_{j,k}(\mathbf{z}) = \frac{\#[t_i=t_j,\ i \in I_k(\mathbf{z})]}{\#[Y_i \geq t_j,\ i \in I_k(\mathbf{z})]},$$

and $I_k(\mathbf{z})$ are the indices of the k nearest neighbors of $\mathbf{z}$.

More generally, suppose $U_n(t|\mathbf{z})$ and $V_n(t|\mathbf{z})$ are strongly uniformly consistent estimators of

$$U(t|z) = P\{Y \geq t|Z=z\} \quad \text{and} \quad V(t|z) = P\{Y \geq t,\ J=1|Z=z\}\ .$$

Beran shows that estimates of the form

$$S_n(t|\mathbf{z}) = \prod_{\{j|t_j<t\}} [1 - \hat{\lambda}_j(\mathbf{z})],$$

with

$$\hat{\lambda}_j(\mathbf{z}) = \left[V_n(t_j|\mathbf{z}) - V_n(t_j^+|\mathbf{z})\right] / U_n(t_j|\mathbf{z})$$

are strongly uniformly consistent over $[0,\tau(\mathbf{z}))$ for any $\tau(\mathbf{z}) < T(\mathbf{z})$, where $T(\mathbf{z}) = \sup\{t : U(t|\mathbf{z}) > 0\}$. In particular, this uniform consistency holds for the BRPLE $S_{n,k}$ provided that $k \to \infty$ and $(k/n) \to 0$ as $n \to \infty$.

One regression type function that measures the influence of the covariates on survival is the median $m(\mathbf{z})$ of the survival function $S(t|\mathbf{z})$. Let

$$\begin{aligned} m^+(\mathbf{z}) &= \inf\{t : S(t|\mathbf{z}) < \tfrac{1}{2}\} \\ m^-(\mathbf{z}) &= \sup\{t : S(t|\mathbf{z}) > \tfrac{1}{2}\} \end{aligned} \qquad (2.1)$$

be the upper and lower medians of $S(t|\mathbf{z})$, respectively. Then the conditional median survival is defined by

$$m(z) = \tfrac{1}{2}[m^+(z) + m^-(z)]. \tag{2.2}$$

Beran shows that if $m_n(z)$ is the estimate obtained by replacing $S(t|z)$ by $S_n(t|z)$ in (2.1) and (2.2), then

$$m_n(z) \xrightarrow{a.s.} m(z)$$

provided $S(t|z)$ has a unique median $m(z)$ satisfying $m(z) \in [0, T(z))$.

We next consider estimates of the regression function

$$\mu(z) = E(T|Z=z) = -\int_0^\infty t\,dS(t|z) = \int_0^\infty S(t|z)dt$$

where we assume that the integrals exist, and, in order to ensure that $\mu(\cdot)$ is identifiable, we assume that

$$\sup\{t:S(t|z)>0\} \le \sup\{t:G(t|z)>0\} \tag{2.3}$$

where $G(t|z) = P(C \ge t|z)$ is the conditional survival function of C. Note that under this assumption, $\sup\{t:S(t|z) > 0\} = T(z)$. A natural estimate of $\mu(\cdot)$ is defined by

$$\mu_n(z) = \int_0^\infty S_n(t|z)dt.$$

Theorem 1. Suppose that $T(z) < \infty$, *and that the preceding assumptions of this section hold; then* $\mu_n(z) \to \mu(z)$ *a.s. as* $n \to \infty$.

Proof. Let $\varepsilon > 0$. Note that

$$|\mu_n(z)-\mu(z)| \le \sup_{0 \le t \le T(z)-\frac{1}{2}\varepsilon} |S_n(t|z)-S(t|z)|\,[T(z)-\tfrac{1}{2}\varepsilon] + \int_{T(z)-\frac{1}{2}\varepsilon}^{T(z)} |S_n(t|z) - S(t|z)|dt.$$

From Beran (1981), we can choose N such that the first term on the right is a.s. bounded by $\frac{1}{2}\varepsilon$ for $n \ge N$. The second term is bounded by $\frac{1}{2}\varepsilon$ for all n and thus the result follows.

2.2 Independent Censoring Variables

Next we turn to the case where the censoring variable C is independent of the covariates Z. We let G denote the survival function $G(t) = P(C \geq t)$. The construction of our estimate is based on the following.

Lemma. If C *is independent of* (T,Z), *if (2.3) holds, and if* $\mu(z)$ *exists, then* $E[JY/G(Y)|Z = z] = \mu(z)$.

Proof. Using our independence assumption, we find

$$E\left[\frac{JY}{G(Y)}\,|Z=z\right] = E\left\{E\left[\frac{JY}{G(Y)}\,|T,\ Z=z\right]|Z=z\right\}.$$

Next, note that

$$E\left(\frac{JY}{G(Y)}\,|T=t,\ Z=z\right) = E\left\{\frac{I[t \leq C](t \wedge C)}{G(t \wedge C)}\,|T=t,\ Z=z\right\}$$

$$= E\left\{t\,\frac{I[t \leq C]}{G(t)}\right\} = \frac{t}{G(t)}\,E\{I[t\leq C]\} = t$$

The lemma follows from this.

Let G_n denote the Kaplan-Meier (1958) product limit estimate of G; the lemma suggests the following estimate of $\mu(z)$:

$$\hat{\mu}(z) = \frac{1}{k} \sum_{i \in I_k(z)} \frac{J_i Y_i}{G_n(Y_i)}.$$

To show the consistency of $\hat{\mu}(z)$, we write

$$\hat{\mu}(z) - \mu(z) = \frac{1}{k} \sum_{i \in I_k(z)} \left[\frac{J_i Y_i}{G(Y_i)} - \mu(z)\right]$$

$$+ \frac{1}{k} \sum_{i \in I_k(z)} J_i Y_i \left[\frac{1}{G_n(Y_i)} - \frac{1}{G(Y_i)}\right] = I + II$$

Using the arguments of Beran, we can show that for any $\tau(z)$ satisfying $\tau(z) < T(z)$,

$$\sup_{0 \le t \le \tau(z)} |G_n(t) - G(t)| \to 0 \quad \text{a.s.}$$

Since

$$\frac{1}{G_n(t)} - \frac{1}{G(t)} = \frac{G(t) - G_n(t)}{G_n(t)G(t)} ,$$

then it follows that

$$\sup_{0 \le t \le \tau(z)} \left| \frac{1}{G_n(t)} - \frac{1}{G(t)} \right| \to 0 \quad \text{a.s.}$$

provided only that $G(\tau(z)) > 0$. Since, by ordinary nearest neighbor regression theory (for example, Collomb (1979)) and the Lemma,

$$\frac{1}{k} \sum_{i \in I_k(z)} \frac{J_i Y_i}{G(Y_i)} \to \mu(z) \quad \text{a.s.} \tag{2.4}$$

we find that $II \to 0$ a.s. uniformly under conditions where (2.4) holds.

Note that I can be rewritten as

$$\frac{1}{k} \sum_{i \in I_k(z)} A_i - \mu(z) ,$$

where $A_i = J_i Y_i / G(Y_i)$ and $E(A_i) = \mu(z)$. Thus the convergence of I can be handled by the theory of ordinary (uncensored) nonparametric regression [for example, Stone (1977) and Collomb (1979)].

Theorem 2. Suppose that C *is independent of* (T,Z), *that* $G(\tau(z)) > 0$, *that the conditions* $\mathcal{D}_2$ *of Collomb holds for* $(A_1,Z_1),\ldots,(A_n,Z_n)$ *and that* $k/n \to 0$, $k/(\log n) \to \infty$. *Then,*

$$\sup_{0\le t\le \tau(z)} |\hat{\mu}(z) - \mu(z)| \to 0 \quad \text{a.s.}$$

Note that by using Stone (1977), we can obtain L^q convergence, $1 < q < \infty$, under much weaker distributional assumptions.

3. WEAK CONVERGENCE

We assume throughout this section that $\mu(z)$ exists, (2.3) holds, $p = 1$, and that C is independent of (T,Z).

In order to establish the asymptotic distribution of $\hat{\mu}(z)$, consider the decomposition

$$\hat{\mu}(z) - \mu(z) - \frac{1}{k}\sum[\mu(Z_i) - \mu(z)] + \frac{1}{k}\sum\left[\frac{Y_iJ_i}{G(Y_i)} - \mu(Z_i)\right]$$

$$+ \frac{1}{k}\sum Y_iJ_i\left[\frac{1}{G_n(Y_i)} - \frac{1}{G(Y_i)}\right] = I + II + III$$

in which the sums are over the set $I_k(z)$. Part I converges to zero a.s. at the rate $(k \log n)^{-\frac{1}{2}}$ [Révész (1979, Lemma 2)]. Földes and Rejtö (1981) showed that

$$\sup_{0\le t\le a} |G_n(t) - G(t)| = O_p((\log n/n)^{\frac{1}{2}}) ,$$

provided $S(\cdot)$ and $G(\cdot)$ are continuous and bounded away from 0 on a finite interval $[0,a]$, where $S(t) = P(T\ge t)$. One can readily show that $\sup|G_n^{-1} - G^{-1}|$ is of the same order. Now, by ordinary nearest neighbor theory, for example Collomb (1979), under appropriate conditions,

$$\frac{1}{k}\sum Y_iJ_i/G(Y_i) \to \mu(z) \quad \text{a.s.}$$

It follows that $III = O_p((\log n/n)^{\frac{1}{2}})$.

For II we need to examine the distribution of $W = YJ/G(Y) - \mu(Z)$. Let $H(w) = P(W\ge w|Z=z)$, where it is assumed that

$$P(W \geq w | Z=z) \text{ does not depend on } z. \tag{3.1}$$

We now have the necessary approximations to prove by analogy the central result of Révész (1979).

Theorem 3. Assume that (Y_i, J_i, Z_i) *are i.i.d. as above, with* $0 \leq Z \leq 1$, *and the density of* Z *bounded away from* 0. *Assume also that* $|\mu'|$ *is uniformly bounded, that (3.1) holds and that*

$$\int_0^\infty e^{xt}\, dH(x) < \infty \quad \textit{for} \quad |t| < t_0 < \infty, \textit{ some } t_0 > 0.$$

Let $k \to \infty$ *as* $n \to \infty$ *in such a way that* $kn^{-2/3}\log n \to 0$, $(\log n)^3/k \to 0$. *Then*

$$P\{k^{\frac{1}{2}}\sigma^{-1} \sup_{0 \leq z \leq 1} |\hat{\mu}(z) - \mu(z)| \leq B(n/k, y)\} \to A(y)$$

where $A(y) = \exp(-2e^{-y/2})$,

$$B(u,v) = (2\log u + \log\log u - \log \pi + v)^{\frac{1}{2}}, \quad u > e,$$

and σ^2 *is the variance of* W.

A natural estimate of σ^2 is

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} (\hat{W}_i - \hat{\mu}(Z_i))^2$$

where $\hat{W}_i = J_i Y_i / G_n(Y_i)$. Note that $\sigma_n^2 = \tilde{\sigma}_n^2 + R_n$ where

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} [W_i - \tilde{\mu}(Z_i)]^2$$

$$\tilde{\mu}(Z_i) = \frac{1}{k}\sum_{i \in I_k(z)} W_i$$

and the remainder term is of order at most the order of $\|G_n - G\|$ provided

$$\frac{1}{k}\sum_{i \in I_k(z)} W_i^2 \to \sigma^2 \quad \text{a.s.},$$

which can be obtained from Collomb (1979). By Földes and Rejtö (1981), $R_n = O_p((\log n/n)^{\frac{1}{2}})$. Thus there exists $N \geq 1$ such that for $n \geq N$,

$$\sqrt{\frac{n}{\log n}}\,|R_n| \leq K \quad \text{a.s.}, \quad K < \infty.$$

So, for $n \geq N$,

$$P\left(\left| \frac{\sigma_n^2}{\sigma^2} - 1 \right| \geq (4\log n/\sqrt{n}) + K\sqrt{\frac{\log n}{\sqrt{n}}} \right) \leq \frac{1}{n^2}$$

by Révész (1979), Lemma 6.

Finally, it should be noted that Révész uses a slightly different nearest neighbor estimate, where $I_k(z)$ is replaced by $\tilde{I}_k(z)$, which is the set of indices giving the $k/2$ nearest neighbors on the left and the $k/2$ nearest neighbors on the right, k even.

With this modification we find

Theorem 4. If the conditions of Theorem 3 hold, if $\hat{\mu}$ and σ_n is defined using $\tilde{I}_k(z)$ instead of $I_k(z)$, then the conclusion of Theorem 3 holds if σ is replaced by σ_n.

4. APPLICATIONS

Nearest neighbor estimators are compared with the proportional hazards (Cox, 1972) and the least squares (Buckley and James, 1979) estimators using data from the Stanford heart transplant program. We examine average (mean or median) survival time after transplant given the covariate of age at transplant. Patients alive beyond the span of available observations, October 1967 to February 1980, were considered as censored. For purposes of comparison with Miller and Halpern (1981), we restrict our attention to the 157 out of 184 cases who had complete tissue typing. The data are displayed in Table A of Miller and Halpern (1981).

During the Stanford heart transplant program it was found that younger patients had better survival after transplant. This in turn may have affected screening of patients in the later years of the study (R. Miller, personal communication). A drop in median age at entry can be seen in Figure 1 as the patient number increases. Thus it appears that censoring time depends on age. The estimators μ_n and m_n used below were derived under this assumption.

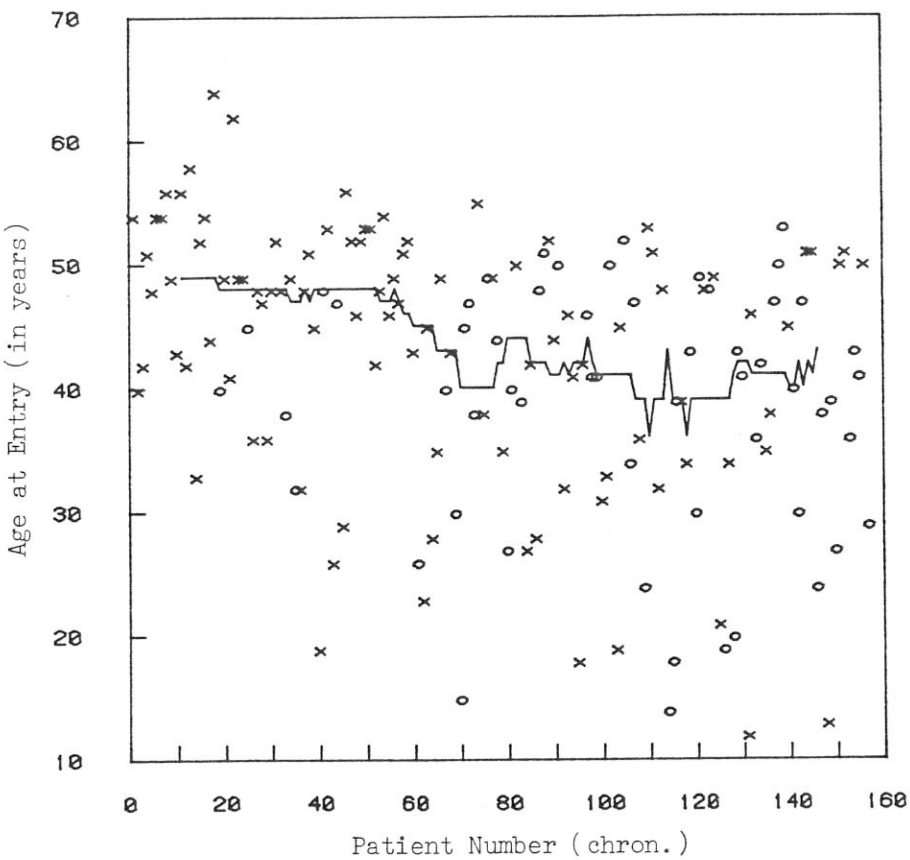

Figure 1. Regression of age at entry on patient number (n = 157). Patients are numbered chronologically from October 1967 to February 1980. x = dead by 2/80; 0 = alive at 2/80. Line is running median (k = 20) regression of patient number on age at entry.

The nearest neighbor regression estimators first considered are the running censored mean $\mu_n(z)$ and the running censored median $m_n(z)$ of Section 2.1. In this section, we used the modified product limit estimate, where the last observed time, whether it is a survival time or censoring time, is treated as a survival time (Efron, 1967). Several choices of k were tried, with $k = 40$ being a happy, but heuristic, medium between a very rough and a very smooth curve. Tied entry ages were included in nearest neighborhoods, bringing the number of neighbors above 40 for many points. This ad hoc solution seemed preferable to selecting a subset of ties, and it washes out in the asymptotics.

The Buckley-James (1979) estimator is drawn directly from Miller and Halpern (1981, Table 1). That is,

$$E(\log T_i | Z_i) = a + Z_i b$$

with $\hat{a} = 3.16$ and $\hat{b} = -0.013$. The median survival time under the Cox (1972) model is found by solving (Miller and Halpern, 1981)

$$\hat{S}_0(\hat{m}_c(z))^{\exp(z\hat{\beta})} = \tfrac{1}{2}$$

for $\hat{m}_c(z)$. Here $\hat{\beta} = 0.028$ and

$$-\log \hat{S}_0(t) = \int_0^t \hat{\lambda}_0(u)du ,$$

with

$$\hat{\lambda}_0(u) = d_i \left[(t_i - t_{i-1})_{\{j \quad R(t_i)\}} \exp(Z_j\hat{\beta}) \right]^{-1}$$

for $t_{i-1} < u < t_i$, in which $t_1 < t_2 < \dots < t_r$ are the ordered distinct uncensored observations and d_i is the number of deaths at t_i. $R(t) = \{j | Y_j \geq t\}$ is the risk set at time t, and Z_j is the covariate for the patient with survival time t_j.

Figure 2 presents the Buckley-James, Cox, and running median ($k = 40$) regression estimates relating survival time to age on log scale. Note that by design the linear Buckley-James estimator must either increase, decrease, or remain constant with age. Similar constraints hold for the Cox model, though it is modulated by the survival time-dependent intensity $\hat{\lambda}_0(t)$. The nonparametric running median has no such constraints. In fact, one sees that this estimator is relatively constant up until age 40. Part of the roughness may be due to the large number of censored observations nearby. At 50 years the estimated survival time drops by an order of magnitude.

Miller and Halpern (1981) considered quadratic curves in addition to the linear curves, restricting attention to the 152

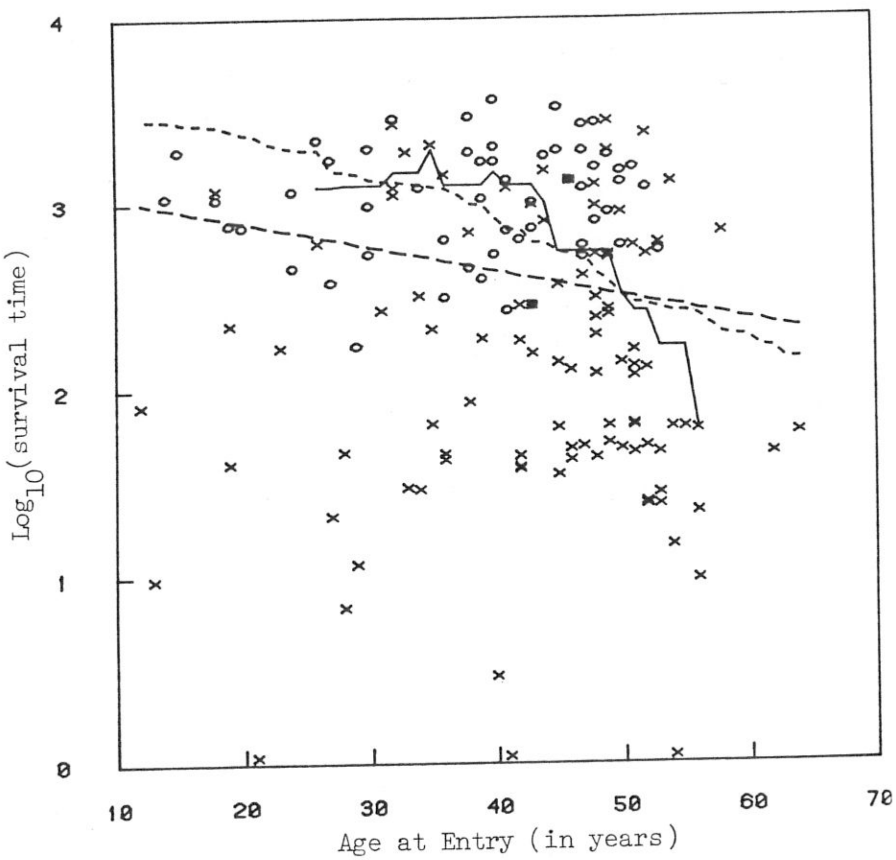

Figure 2. Regression of survival time on age in $\log_{10}$ scale (n = 157). x = dead by 2/80; 0 = alive at 2/80. Solid line = running median (k = 40). Short dash = Cox with linear term. Long dash = Buckley-James with linear term.

cases who survived at least 10 days. The Buckley-James curve

$$\hat{E}(\log T_i | Z_i) - 1.21 + 0.113 Z_i - 0.0017 Z_i^2$$

and the estimated median survival time from the Cox model, with

$$(Z_i, Z_i^2)\beta' = -0.149 Z_i + 0.0024 Z_i^2 ,$$

are presented in Figure 3 along with the running median (BRPLE) (k = 40). The initial lower survival for the Cox and Buckley-James is an artifact of the quadratic fit. However, the nearest neighbor curve does not extend below age 26 because of the edge effect.

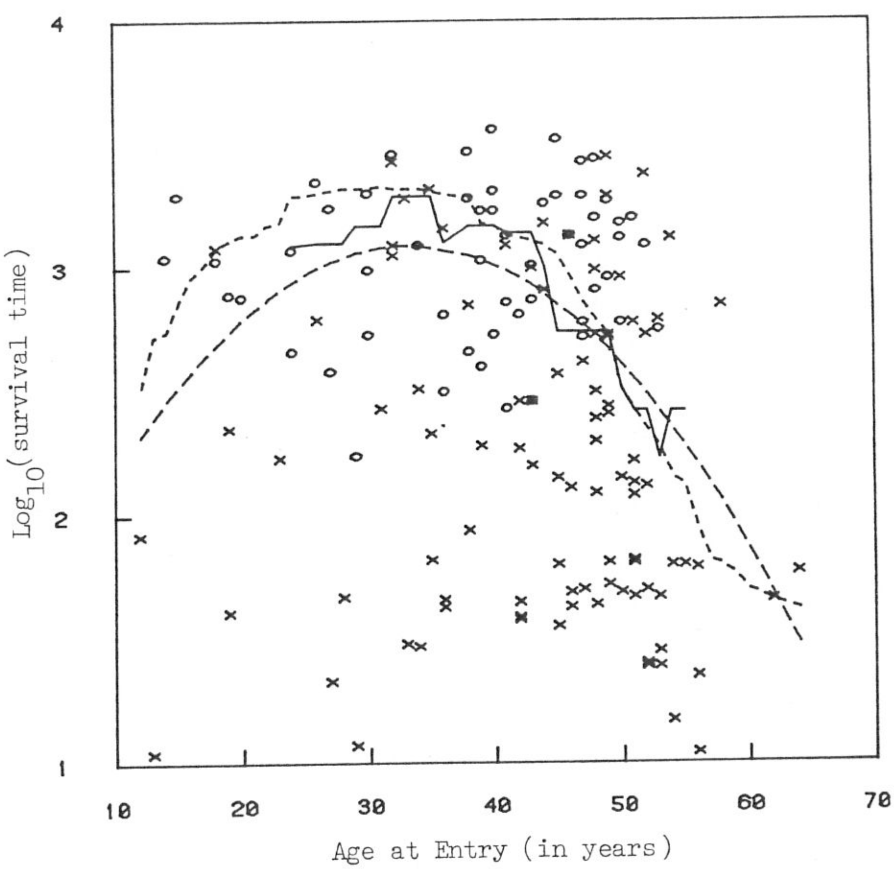

Figure 3. Regression of survival time on age in $\log_{10}$ scale (n = 152). See Figure 2 for symbols. Cox and Buckley-James have quadratic terms.

The running mean regression (k = 40) is presented in Figure 4 along with the running median (BRPLE). It is higher than the running median because at each z the conditional survival distribution is skewed to the right. Note that the running mean does not display as great a drop in the higher age category. This is probably due to the few large uncensored observations, which outweigh the many small survival times. The third curve is the running mean of $\log_{10}$ (survival time).

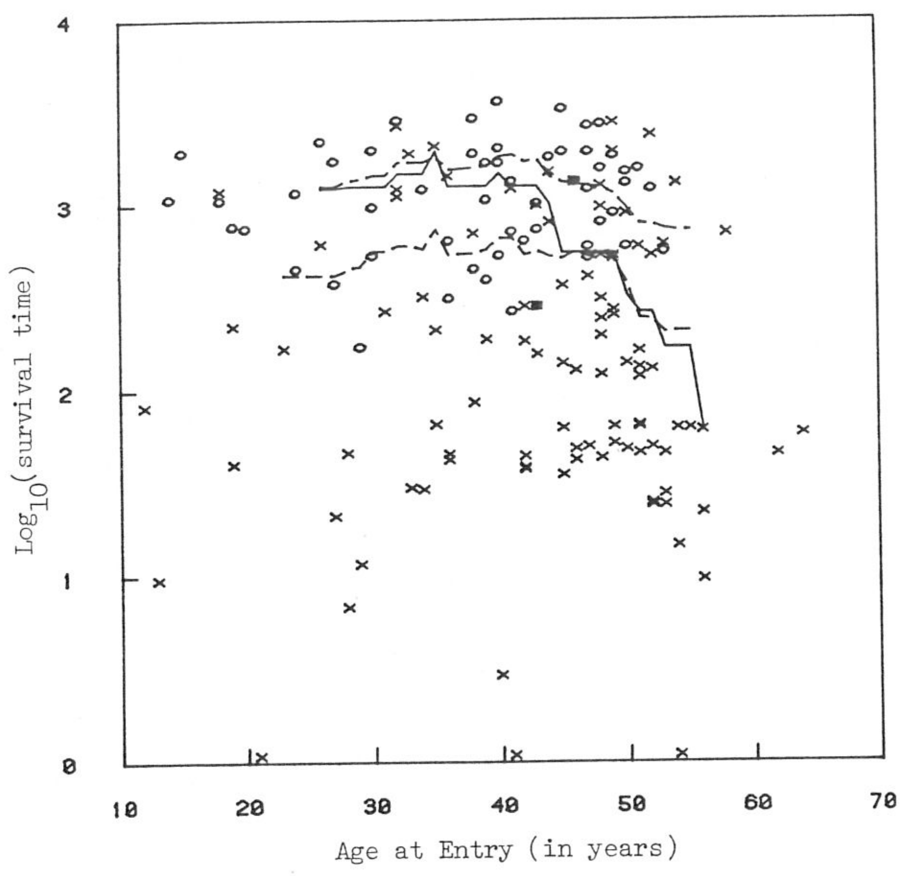

Figure 4. Regression of survival time on age in $\log_{10}$ scale (n = 157). Running median (solid line), running mean (dot-dashed line), and running mean of $\log_{10}$ (survival time) (dashed line) with k = 40.

Although the censoring distribution seems to depend on age, one might hope that this effect is minimal. Figure 5 presents the 40 nearest neighbor running mean, $\hat{\mu}(z)$, and median curves under the assumption that censoring does not depend on the covariate age.

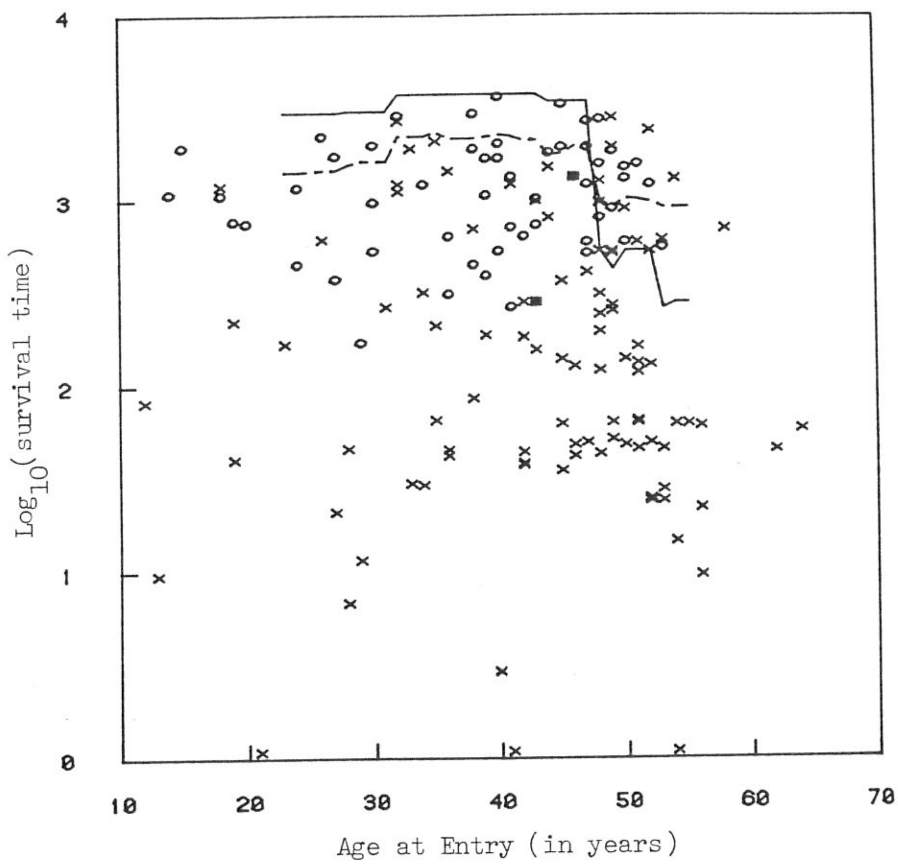

Figure 5. Regression of survival time on age in $\log_{10}$ scale ($n = 157$). Censoring assumed independent of covariates.

Here the median is the median of the "empirical" distribution giving weight

$$\frac{J_i}{G_n(Y_i)} \Big/ \sum_{i \in I_k(z)} \frac{J_i}{G_n(Y_i)}$$

to the points Y_i, $i \in I_k(z)$. The median is so high because of the large censored survival times.

Note added in proof: We recently learned that there were two transcription errors in the Stanford heart transplant data used by Miller and Halpern in their April 81 Stanford Technical Report No. 66 and by us in this paper. The corrected values (for patients 66 and 127) are given by Miller and Halpern (1982). The changes in the data produce only minor changes in the estimates, and all the estimates given in this paper are computed for the same data set.

REFERENCES

BERAN, R. (1981), "Nonparametric Regression with Randomly Censored Survival Data," Unpublished manuscript, Univ. of California, Berkeley.

BUCKLEY, J., and JAMES, I. (1979), "Linear Regression with Censored Data," *Biometrika*, 66, 89-99.

COLLOMB, G., (1979), "Estimation de la Regression par la Methode des k Points les Plus Proches avec Noyau: Quelques Propriétés de Convergence Ponctuelle," Statístique Non Parametrique Asymptotique, Rouen, June 1979, *Springer Lecture Notes in Math.*, 821, 159-175.

COX, D.R. (1972), "Regression Models and Life Tables," *Journal of the Royal Statistical Society*, Ser. B, 34, 187-202.

——— (1975), "Partial Likelihood," *Biometrika*, 62, 269-276.

EFRON, B. (1967), "The Two Sample Problem with Censored Data," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability IV*, University of California, Berkeley, 831-853.

FÖLDES, A., and REJTÖ, L. (1981), "Strong Uniform Consistency for Nonparametric Survival Curve Estimators from Randomly Censored Data," *Annals of Statistics*, 9, 122-129.

KAPLAN, E.L., and MEIER, P. (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53, 457-481.

KOUL, H., SUSARLA, V., and VAN RYZIN, J. (1981), "Regression Analysis with Randomly Right Censored Data," *Annals of Statistics*, 9, 1276-1288.

LEHMANN, E.L. (1953), "The Power of Rank Tests," *Annals of Mathematical Statistics*, 24, 23-43.

MILLER, R.G. (1976), "Least Squares Regression with Censored Data," *Biometrika*, 63, 449-464.

MILLER, R., and HALPERN, J. (1982), "Regression with Censored Data," *Biometrika*, 69, No. 3.

RÉVÉSZ, P. (1979), "On the Nonparametric Estimation of the Regression Function," *Problems of Control and Information Theory*, 8, 297-302.

STONE, C.J. (1977), "Consistent Nonparametric Regression," *Annals of Statistics*, 5, 595-645.